PIETRAGALLO
PIETRAGALLO GORDON ALFANO
BOSICK & RASPANTI, LLP
ATTORNEYS AT LAW

38TH FLOOR ONE OXFORD CENTRE PITTSBURGH, PA 15219
412.263.2000 FAX: 412.263.2001
WWW.PIETRAGALLO.COM

DIRECT DIAL NO.: 412.263.1816
DIRECT FAX NO: 412.263.4246
FILE NO.: MYLAN-112578
EMAIL: cct@pietragallo.com

May 27, 2020

Via ECF

The Honorable Joel Schneider
United States Magistrate Judge
USDC, DISTRICT OF NEW JERSEY
Mitchell H. Cohen Building & U.S. Courthouse
4th & Cooper Streets
Camden, NJ 08101

Re: *In re Valsartan, Losartan, and Irbesartan Products Liability Litigation*
USDC, District of New Jersey, No. 1:19-md-2875-RBK-JS

Dear Judge Schneider:

The ESI search terms incorporated in the Court's December 23, 2019 Order, Dkt. 328, are demonstrably overbroad and will impose a disproportionate burden on the Manufacturer Defendants absent significant modification. At the December 11, 2019 case management conference, the Court reemphasized what it had said "so many times before" – that "***[w]e're going to set the search terms, the custodians, the document requests, but if there is good cause on either side to either add or subtract from what we agree to and order, the Court will entertain the application***." 12/11/19 CMC Tr. 26:17–23 (emphasis supplied).

There is now good cause for the Court to adopt the modifications proposed by the Manufacturer Defendants. For example, Mylan's counsel conducted a sample review of documents captured by the ESI search terms, and determined that ninety-four percent (94%) of the documents reviewed were non-responsive, and less than two percent (2%) of the responsive documents were "true" hits. Similarly, counsel for Teva conducted a sample review of returned results from two separate documents sets, one set captured by the original terms and a second set captured using Defendants' counter proposal terms. The review revealed that while 14.8% of documents sampled from the counterproposal hit reports were responsive, only 1.3% of documents in the sampled set utilizing the original search terms were responsive. This low prevalence of responsive documents clearly demonstrates that the ESI search terms need to be revised.

Thus, the Court should adopt the Manufacturer Defendants' proposal, which will increase the efficiency of discovery, will drastically reduce waste, and will lessen the significant burden on the Manufacturer Defendants.[1]

[1] As explained in more detail below, Zhejiang Huahai Pharmaceutical Co, Ltd., Solco Healthcare U.S., LLC, Prinston Pharmaceutical Inc., and Huahai U.S., Inc. (the "ZHP Parties") believe many

I. Plaintiffs have refused to engage in the process to which they previously agreed.

The use of Boolean search terms and related methodologies during discovery to capture ESI is an inherently iterative process. *See Walker v. N. Las Vegas Police Dep't*, No. 2:14-cv-1475, 2016 WL 8732300, at *3 (D. Nev. May 13, 2016) (stating, "[w]here, as here, counsel are using keyword searches for retrieval of ESI, the parties must work together to implement an iterative process for finalizing the search terms." (internal citation, quotation, and modifications omitted)). Early on in this litigation, the parties recognized and accounted for the iterative nature of the process in the ESI Protocol, which was entered by this Court nearly a year ago. Dkt. 127. In addition to requiring that the parties meet and confer to develop initial search terms and related protocols, the ESI Protocol expressly contemplates that the process will be ongoing throughout this litigation:

> The parties will continue to meet and confer in accordance with the Federal Rules of Civil Procedure and Local Civil Rules of this court regarding any search process issues as necessary and appropriate. Nothing in this protocol, or the subsequent designation of any search terms, shall operate to . . . limit the ability to agree to meet and confer to modify the search terms, criteria, or methodologies ***as information is obtained and utilized in the course of the litigation***.

Dkt. 127, ESI Protocol, at 4 (emphasis supplied).

The parties subsequently engaged in a meet-and-confer process regarding ESI search terms, which was conducted in parallel with negotiations regarding ESI custodians and Plaintiffs' Rule 34 document requests. Indeed, ESI custodians and Plaintiffs' Rule 34 document requests – many of which were narrowed or eliminated by the Court – were not finalized until December 23, 2019, via the same Order that incorporated the ESI search terms that are now in dispute. *See* Dkt. 303. Thus, due to the timing of the process, the Manufacturer Defendants were unable to obtain initial hit counts or to perform document sampling before the Court approved the search terms in December 2019. *See, e.g.*, *Arconic Inc. v. Novelis Inc.*, No. 2:17-cv-1434, 2019 WL 5802365, at *19 n.9 (W.D. Pa. Sept. 6, 2019) (determining that the collection, processing, and searching of the data of merely proffered custodians would be an extraordinarily expensive and inefficient way to determine if a proposed custodian's data is cumulative of other custodians). The Manufacturer Defendants were likewise unable to specifically demonstrate the prevalence of false positives, nor were they in a position to appreciate the associated financial burden, without a complete understanding of how the requests for production, search terms, and custodians would coalesce.

---

of the issues addressed in this motion are not yet ripe as to the ZHP Parties because they are not in the same position as the other Manufacturer Defendants. As previously explained to and acknowledged by the Court, the COVID-19 pandemic caused a substantial delay in collecting ZHP's data in China. *See* March 16, 2020 Order, Dkt. 400 ("The parties have already been informed the Court will make all reasonable accommodations in view of the virus situation.").

Even so, the Manufacturer Defendants made clear to Plaintiffs and the Court that the contemplated search terms were facially overbroad and would lead to wasteful discovery. *See* Dkt. 312, Defs.' 12/5/19 Letter Br., at 18-31. Plaintiffs, however, would not agree to further compromise until the Manufacturer Defendants provided hit counts – which, for reasons already explained, the Defendants could not provide. Thus, in order to break the impasse and move the process forward, Defendants acquiesced to the search terms attached to the Court's December 23, 2019 Order (hereafter, the "Original Search Terms"), but only on the condition that the Original Search Terms could be validated and continuously refined as information became available through test searches and targeted review. On December 10, 2019, Plaintiffs wrote to the Court memorializing the parties' agreement with regard to that methodology:

> Defendants believe that many of these terms and modifiers, as proposed, will require substantial unnecessary efforts and cost. ***The Parties continue to be willing to refine and narrow these search terms in order to reduce unnecessary efforts and cost*** while also minimizing the risk that relevant documents will be excluded and therefore not be reviewed for responsiveness to Plaintiffs' document requests. To that end, ***the Parties have devised and agreed upon the detailed method described below for further evaluating disputed search terms*** while simultaneously reviewing and producing non-custodial documents and documents resulting from undisputed search terms. The Parties respectfully submit that this procedure will allow for the most expeditious rolling production of documents by Defendants ***while giving the Parties the ability to further refine the search terms as necessary based upon actually collected documents and data.***

Dkt. 316, Pls.' 12/10/19 Letter Br., at 7–8 (emphasis added).

The "detailed method" referenced by Plaintiffs was memorialized in the December 23, 2019 Order. It included the ability for Defendants to narrow broad categories of the Original Search Terms, including Manufacturing terms, Medical Condition terms, Economic terms, Entities terms, QA-Testing terms, cGMP terms, and Regulatory terms. The terms within these categories were to be tested and sampled by the Manufacturer Defendants and, if the parties could not agree to modifications following a meet and confer, the terms were to be brought to the Court for resolution. Dkt. 328, Or., at 56–57.

In addition to Defendants' ability to refine, on a wholesale basis, multiple categories of the Original Search Terms based on testing and sampling, the parties also delineated thirty-one (31) individual search terms that were objected to in their entirety by the Defendants. With regard to these terms, which were highlighted in yellow on the Original Search Terms spreadsheet submitted to the Court, the parties contemplated that Defendants would run hit counts and, assuming as Defendants suspected that the terms were overly broad, the parties would modify these terms without the need for any additional testing and sampling. *Id.* at 57.

A. Mylan

Upon the entry of the December 23, 2019 Order attaching the approved Rule 34 requests, the custodian lists, and the Original Search Terms, the Manufacturer Defendants began the data collection process. Mylan, for example, began its initial collection of custodial Microsoft Outlook email files on January 2, 2020. (Ex. A, Mylan Decl. ¶ 3.) Less than two weeks later, Mylan had compiled an initial data set and generated hit counts based on the Original Search Terms. (*Id.*, ¶ 4.) Two things were immediately apparent. First, the total volume of hits being generated – more than 30 percent (30%) of all emails sent or received by Mylan's custodians over nearly a decade – was far greater than the parties had anticipated. And, second, a few readily identifiable issues[2] were isolated as sources of false positives. With regard to the second issue, counsel for Mylan reached out to the Plaintiffs in order to set up an initial meet and confer to address certain easily resolvable issues (Ex. B, 1/16/20 Email from F. Stoy to B. Parekh.) Plaintiffs did not immediately respond to this request.

In the meantime, as Mylan continued its analysis, it became apparent that incremental changes to the Original Search Terms would not be sufficient to materially decrease the extremely high volume of data being captured. Thus, on January 20, 2020, counsel for Mylan followed up with Plaintiffs renewing the request to meet and confer and advising that Mylan's "custodial document collection and review is on hold pending the resolution of several threshold issues related to search terms." (Ex. B, 1/20/20 Email from F. Stoy to B. Parekh.) Indeed, Mylan could not even submit its custodial Microsoft Outlook email files to its eDiscovery vendor for de-duplication, threading, and further analysis because, due to the extreme volume being captured, it would cost Mylan approximately $2 million just to process the data. (Ex. C, Consilio Decl. ¶ 6.)

Mylan and Plaintiffs conducted an initial meet and confer the next day, and Plaintiffs agreed to modify two primary search terms ("destroy" and "DMF") in order to eliminate obvious false positives. During that discussion, counsel for Mylan made Plaintiffs aware of the need for additional negotiations in light of the total volume of data being captured by the Original Search Terms, including several search terms that were each yielding over 1,000,000 email hits. Counsel for Mylan further advised that additional testing and analysis would be performed to facilitate a substantive discussion.

Thereafter, Mylan continued to collect data and test the Original Search Terms, making modifications beginning with the thirty-one (31) "highlighted terms," and then expanding to broader swaths of search terms when piecemeal changes to individual terms did not reduce the total volume of data due to the large amount of overlap between the hundreds of Original Search Terms approved by the Court. Through this labor-intensive process, Mylan determined that significant modification to the Original Search Terms was required.

---

[2] By way of illustration, one of the issues initially identified by Mylan was that the search term "destroy" was contained in the email footer for all Mylan custodians (notifying unintended recipients to "immediately destroy all electronic, paper and other versions" of the message), thus leading to the inclusion of every email containing that standard disclaimer in the document collection.

On March 19, 2019, Mylan submitted its counterproposal to Plaintiffs (the "Revised Search Terms").[3] The Revised Search Terms were designed to reduce both the total volume of hits and the number of false positives without losing sight of Plaintiffs' purpose in including a particular term among the Original Search Terms. Over the next several weeks, Mylan continued to share additional information with Plaintiffs, including detailed hit counts for both the Original Search Terms and the Revised Search Terms, a redline comparison between the two proposals, explanations regarding the methodology used to develop the Revised Search Terms, and the results of a sample set review demonstrating the overwhelming number of false positives being generated by the Original Search Terms.

Plaintiffs have nonetheless refused to engage in the meet-and-confer process that was agreed upon by the parties and adopted by the Court in its December 23, 2019 Order. Plaintiffs' refusal has placed Mylan in an untenable position where it must incur enormous costs in order to process, store, and review documents that are largely non-responsive to Plaintiffs' document requests and irrelevant to this litigation. With no other avenue of relief, Mylan has been left with no choice but to seek redress from the Court.

B. Teva

The Teva Defendants experienced similar difficulties with the search terms. In January 2020, the Teva Defendants initiated collection of custodial files from the 36 custodians identified by Teva. Given the huge volume of documents implicated by this large number of custodians, and the anticipated impact both in terms of time and expense of processing and searching across this dataset, on January 31, 2020, the Teva Defendants identified a subset of 16 key custodians across all business functions to serve as a sample group for testing the original agreed upon search terms. (Ex. E, Teva Decl. ¶ 2.) The Teva Defendants' vendor collected and processed the complete custodial email files for these 16 custodians. (*Id.* ¶ 3.) Due to the breadth and complexity of the agreed upon search terms, the Teva Defendants spent the next ten days working to mitigate issues that arose from applying these terms as drafted across the processed custodial files. (*Id.*). The Teva Defendants received their first detailed hit reports on February 14, 2020, which revealed that these search terms would – after deduplication - require promotion for review of approximately 52.3% of the entire dataset for these 16 custodians. (*Id.* ¶ 4.) Over the following week, the Teva Defendants met with their vendor on multiple occasions to analyze the data and develop methods to decrease the exceedingly high volume of hit counts. Thereafter the vendor generated three more versions of hit count reports as the Teva Defendants worked to identify and remediate terms potentially generating false positives. (*Id.* ¶ 5.) Based on discussions with the vendor, the Teva Defendants reached out to meet and confer on discrete issues identified in the hope that these and similar modifications would lead to a substantial reduction in the overall hit count. (Ex. H, 2/22/20 Email from S. Harkins to B. Parekh.)

[3] On April 24, 2020 the Manufacturer Defendants jointly adopted the Revised Search Terms for purposes of negotiating with the Plaintiffs regarding necessary global revisions. (Ex. D.)

The parties held a productive meet and confer on February 24, 2020. Plaintiffs' counsel expressed approval for a number of Defendants' proposed limiting measures, which the Teva Defendants then implemented in a revised search. (*Id.*). On March 2, 2020, the Teva Defendants approved additional changes suggested by the vendor for eliminating likely junk files in the hit reports. (*Id.*) The Teva Defendants received another set of revised hit reports on March 12, 2020. (*Id.*) However, it quickly became clear that the numerous combinations of terms in the original protocol made incremental changes and modifications largely fruitless. Despite modifications both approved by Plaintiffs, such as eliminating hits based on language including "delete" or "destroy" in standard email disclaimers, and efforts by the vendor such as identifying and excluding junk files being captured in document families, the overall volume of documents requiring promotion for review was still--after deduplication--44.6% of all documents in the custodial files. (*Id.* ¶ 7.) These results indicated that the search terms as originally drafted were not functioning to capture relevant documents, but instead were capturing huge numbers of irrelevant documents.

The following week, Mylan submitted its counterproposal to Plaintiffs. (*See supra*.) The Teva Defendants informed Plaintiffs and the Court of the significant overbreadth of the search terms at the very next hearing, the April 15th Teleconference. (04.15.2020 Trans. at 16-17.) The Teva Defendants again raised the overbreadth of the search terms ahead of the April CMC. (Dkt. 419 at 5). At this time, rather than conduct parallel negotiations or continue to attempt incremental modifications which had not altered the huge overbreadth of hits, the Teva Defendants tested Mylan's counterproposal, with company specific modifications as needed, and saw the first substantial reduction in the overall number of hits. (Ex. E, Teva Decl. ¶ 8.) Notably, however, even using Mylan's counterproposal terms still results in 30.8% of all documents in the relevant custodial files being promoted for review after deduplication. (*Id.*) The Teva Defendants thereafter joined the other Defendants in supporting adoption of Mylan's counterproposal as a joint defense proposal for modification of the search terms.

C. ZHP

While the ZHP Parties were not on the same meet and confer track as Mylan and Teva due to the timing of their document collection in light of COVID-19, Plaintiffs have also demonstrated an unwillingness to engage in the agreed upon meet and confer process with the ZHP Parties. After the finalization of the English language search terms, counsel for the ZHP Parties and Plaintiffs negotiated appropriate Chinese translations for those terms, ultimately coming to an agreement on February 24, 2020. However, by that time, there was significant disruption to the ZHP Parties' operations due to the pandemic, and significant travel restrictions that prevented the ZHP Parties' document vendor from collecting documents in China, as counsel for the ZHP Parties previewed for the Court in a January 31, 2020 email (informing the Court that the Chinese Government had extended the Lunar New Year holiday and was preventing employees from returning to work, and that ZHP's vendor could not travel to ZHP facilities to collect data), and as discussed during the conferences with the Court on February 26, 2020 (2/26/20 Tr. 29:9-30:16), March 11, 2020 (3/11/20 Tr. 18:23-19:13), and April 15, 2020 (4/15/20 Tr. 14:13-15:25). Despite this delay, the ZHP Parties made significant progress in late-April and May in collecting data in both China and the U.S. To date, the ZHP Parties have collected approximately 7.5 terabytes of data, and have

been able to process, de-duplicate, and thread approximately a quarter of a terabyte of data that they have collected from their U.S.-based custodians.

While the ZHP Parties' collection of ESI in China is now complete, the approximately 7.25 terabytes of data collected in China is still being processed (including performing de-duplication and email threading) and searched. Consequently, although the ZHP Parties have been able to begin the meet and confer process with Plaintiffs regarding data collected from its U.S.-based entities, they have only been able to raise very preliminary issues with Plaintiffs regarding the data collected in China with Plaintiffs. Although the ZHP Parties have discussed with Plaintiffs the fact that they will need more time to process this data in order to engage in a meaningful meet and confer process on the vast majority of their data, Plaintiffs intend to render those discussions meaningless by indicating that they will object to raising issues with the Court that arise from such discussions *beyond the briefing schedule for the instant motion*, an artificial deadline that has no basis in any order of this Court or any of the Federal Rules of Civil Procedure. Moreover, Plaintiffs' artificial deadline is directly contrary to the agreed-upon meet and confer procedure set forth in the December 23, 2019 Order, Dkt. 328 at 55-57, described above, and the ESI Protocol.

Notably, Plaintiffs are refusing to engage Mylan in the agreed upon meet and confer process because its searches were purportedly "run without any attempt at determining unique hits (that is, the number of hits that did not already hit another term), and prior to any de-duplication and e-mail threading having been performed . . .", Dkt. 430-1, 5/12/20 Letter from A. Slater to C. Trischler, while at the same time denying the ZHP Parties any meaningful opportunity to engage in the meet and confer process once the ZHP Parties' data is de-duplicated and threaded. Plaintiffs are thus effectively using the delays caused by COVID-19 to deny the ZHP Parties the opportunity to meet and confer. Given that the previously acknowledged situation in China constituted "good cause" for the ZHP Parties' delay in collecting information earlier this year, the ZHP Parties were simply expecting to follow the original plan agreed upon in December and memorialized in the December 23, 2019 Order to test search terms, negotiate with the Plaintiffs, and raise issues with the Court if needed, and the ZHP Parties ask that they be allowed to do so.

Once the ZHP Parties' document review is under way, the ZHP Parties will be able to conduct the type of testing and sampling that Mylan has been able to conduct. As Plaintiffs very well understand, such testing, sampling, and further negotiation with Plaintiffs will not delay the ZHP Parties' review and production of documents. The testing, sampling, and negotiation will in fact be concurrent with the ZHP Parties' document review. Indeed, this is the normal, on-going process and this is precisely why the Parties agreed to it in December.

II. Defendants have demonstrated good cause to modify the Original Search Terms.

The Court has repeatedly assured the parties that any discovery order – including the December 23, 2019 Order regarding the Original Search Terms – will be revised for good cause. *See, e.g.*, 12/11/19 CMC Tr. 26:17–23. Based on the information presented herein by the Manufacturer Defendants, which is the product of hundreds of hours of analysis, the Manufacturer Defendants have demonstrated that good cause exists to modify the Original Search Terms. And

in the absence of any meaningful counterproposal by Plaintiffs, this Court should adopt the Revised Search Terms in full.

The rules of discovery do not entitle Plaintiffs to obtain every document in Defendants' possession that potentially relates to the claims and defenses at issue. *See City of Sterling Heights Gen. Emples. Ret. Sys. v. Prudential Fin., Inc.*, No. 2:12-cv-5275, 2015 WL 5055241, at *2 (D.N.J. Aug. 21, 2015) ("While the right to discovery is thus broad, it is not unlimited, and courts may in the exercise of their discretion deny 'unreasonably cumulative' discovery requests."). Indeed, "there is no obligation on the part of a responding party to examine every scrap of paper in its potentially voluminous files, and in an era where vast amounts of electronic information is available for review, courts cannot and do not expect that any party can meet a standard of perfection." *Enslin v. Coca-Cola Co.*, No. 2:14-cv-6476, 2016 WL 7042206, at *3 (E.D. Pa. June 8, 2016) (internal citation, quotation, and modification omitted); *see also NXP B.V. v. Blackberry, Ltd.*, No. 6:12-cv-498, 2013 WL 6768350, at *5 (M.D. Fla. Dec. 20, 2013) ("[T]he Federal Rules of Civil Procedure do not require perfection or a guarantee that every possible responsive document has been found or produced. When a party responds to discovery requests the Rules require only that a search for responsive materials be objectively reasonable.").[4]

Simply put, "the Federal Rules of Civil Procedure require only a reasonable search for responsive information pursuant to a 'reasonably comprehensive search strategy.'" *Enslin*, 2016 WL 7042206, at *3 (quoting *Treppel v. Biovail Corp.*, 233 F.R.D. 363, 374 (S.D.N.Y. 2006)). Moreover, because the party that is obligated to produce responsive documents typically has superior insight regarding its own data, Court's routinely allow the producing party to define the contours of the search strategy, including reasonable search terms. *See, e.g.*, *id.* at *14 (recognizing that the producing party is "best situated to evaluate the procedures, methodologies, and technologies appropriate for . . . producing [their] own electronically stored information"); *Firefighters' Ret. Sys. v. Citco Grp. Ltd.*, No. 3:13-cv-373, 2018 WL 276941, at *4 (M.D. La. Jan. 3, 2018) (referencing the generally accepted presumption that the producing party "is in the best position to choose an appropriate method of searching and culling data"). This Court has similarly acknowledged that Plaintiffs are not entitled to "free reign" over discovery, and that the Court "must reasonably limit plaintiffs' discovery in order to prevent duplicative, cumulative, and minimally important discovery." 11/20/19 CMC Tr., at 8:12–14, 9:16–17.

Good cause for adoption of the Revised Search Terms exists here because (i) the Original Search Terms are resulting in a massive volume of "hits," (ii) based on document sampling, the vast majority of documents captured by the Original Search Terms are false positives, and (iii) the immense cost for processing, storing, and reviewing documents captured by the Original Search Terms is not proportionate to the needs of this litigation.

---

[4] *See also* THE SEDONA CONFERENCE, *The Sedona Conference Best Practices Commentary on the Use of Search & Information Retrieval Methods in E-Discovery A Project of the Sedona Conference Working Group on Electronic Document Retention & Production (Wg1)*, 15 Sedona Conf. J. 217, 222 (2014) (collecting cases).

A. The Original Search Terms result in astronomical hit counts.

The threshold problem with the Original Search Terms is their facial overbreadth. The Manufacturer Defendants were concerned about this when the Original Search Terms were being negotiated, which was the principal reason why the parties adopted an agreed-upon methodology to further refine and revise the search terms on an ongoing basis as additional information was developed. Using Mylan as an example, as of May 1, 2020, there were 18,540,599 (5.63 TB) of total emails in the Microsoft Outlook email population for Mylan's fifty-two (52) ESI custodians. (Ex. A, Mylan Decl. ¶ 5.) When run against this population, the Original Search Terms yielded 5,718,381 (3.93 TB) email hits. This amounts to approximately thirty-one percent (31%) of the total custodial email population. (*Id.* ¶ 6.) Once attachments to emails are accounted for, the total population, following de-duplication and threading, that would be subject to review could be in excess of ***7,000,000 individual documents***. (*Id.* ¶¶ 7-8.)

This astronomical volume is partially explained by the fact that the primary terms included in the Original Search Terms read like a composite of common words and phrases in the pharmaceutical industry. While this is problematic in its own right, the problem is compounded by the haphazard manner in which modifier terms are applied to primary terms. Typically, modifier terms would serve to limit results, but here the modifier terms do little to refine hits because, as illustrated below, ***entire categories*** of primary terms are run with ***entire categories*** **of** modifier terms. Further, the Original Search Terms largely omit the use of proximity limitations. This is a powerful tool that is typically used to ensure that documents captured for review contain a primary term and a modifier term close enough in proximity as to be related. Without proximity limitations, random combinations of terms will yield hits that do not reflect any association between the primary and modifier terms.

These threshold deficiencies are best demonstrated by example. The Original Search Terms include a category of primary terms related to "QA-Testing." That category consists of fifty (50) primary terms that are highly generic given that the Manufacturer Defendants all engage in quality assurance testing related to each and every one of the medications they manufacturer. Thus, one would expect these primary search terms to be modified in a manner that would limit the results to the drug (valsartan) or impurities (nitrosamines) at issue. Instead, each and every primary term in the QA-Testing category is run using the following modifier script:

> *<QA-Testing Term> AND (<Drug Name> OR (<Solvents> AND NOT <other drug names>) OR <manufacturing modifiers>).*

When built out, the syntax of this script reads as follows:

> *<QA-Testing Term> AND ((*valsartan* or amlodipine* or Diovan* or Exforg* or HCT or HCTZ or sartan or sartans) OR ((*azide* or *formaldehyde* or *-N3* or *xylene* or acid* or aqueous* or chloramin* or HNO2 or hydrochlor* or N-3* or NaNO* or tolue* or Tributyl* or TIN or Triethyl* or TEA or zinc /3 chloride or ZnCI*) AND NOT <other drug names>) OR ((backup/3 data or (impuri* and degradant*) or recall* or residu* or*

solvent or spike* or test* or (unknown /3 peak) or (unknown /3 spike) or (unkown /3 impurit*) or (unknown /3 contamin*))))*.

When Mylan ran the foregoing search related to the QA-Testing terms, the search generated the following hit counts.

	Term	Item Count	GB volume
2	"anion-exchange" or anionexchange or "anion exchange"	4,060	13.93
3	"ion-exchange" or ionexchange or "ion exchange"	11,906	37.99
4	*PCRC*	566	4.04
5	aberran*	70,915	300.59
6	abnorm*	182,813	393.43
7	ALS	20,164	130.31
8	bioequiv*	128,416	402.04
9	CAPA or corrective pre/5 action	194,552	528.25
10	Chromato*	522,652	1,054.72
11	complain*	380,284	771.89
12	contami*	279,230	725.91
13	degradant* or impurit*	1,061,826	1,556.48
14	detect*	517,464	1,146.88
15	EI-MS or EL-MS or "electron ionization"	271	1.23
16	elude*	1,517	1.82
17	fail*	782,683	1,290.24
18	fatal*	46,698	122.58
19	GC or "GC-FID" or GCMS or "GC-MS"	396,263	741.72
20	hazard*	136,960	356.71
21	headspace	60,871	259.68
22	heat	187,544	550.89
23	HPLC-UV	6	0.04
24	HS	72,926	230.31
25	incomplet* /5 data*	16,666	58.84
26	LCMS	60,330	137.75
27	mass spectro*	46,547	128.00
28	method /5 qualification	19,364	59.08
29	noise*	69,780	184.76
30	noti* w/5 pharm*	48,331	224.12
31	obscure*	8,112	19.81
32	observation*	533,343	1,157.12
33	peak*	466,662	1,116.16
34	press /4 release	54,044	105.24
35	preventative /5 action	26,038	92.84
36	problem*	461,156	762.61
37	puri*	647,728	1,228.80
38	quality	1,959,713	2,211.84
39	recall*	446,776	608.17
40	repe* /4 error*	10,122	78.69
41	residu*	577,434	1,075.20
42	retrospec*	72,264	195.03
43	risk*	795,145	1,361.92
44	signal*	109,350	293.18
45	spectro*	161,773	511.86
46	spike*	131,248	364.61
47	suitab*	589,817	1,198.08
48	toxic*	209,330	461.70
49	unknown* or unk or unks	450,168	1,024.00
50	validat*	1,290,415	1,873.92

Clearly, the QA-Testing terms—many of which were highlighted in the Original Search Terms and were therefore explicitly recognized to be problematic—require substantial revision. In refining the search terms, Mylan used the same general format of the Original Search Terms as a template, but rather than making haphazard changes, Mylan analyzed each primary term in order to determine which modifier terms were logically related and would be more likely to yield hits responsive to Plaintiffs' document requests. In contrast, the original search term framework contemplated running entire categories of terms against entire categories of modifiers. Not only was this methodology difficult to execute from a data collection standpoint given the volume of custodians and data, it was not yielding targeted results. The terms on the regulatory, cGMP, and QA Testing pages were modified consistent with this methodology. Mylan also removed certain terms (such as: fail*, observation, problem, quality, headspace, unknown* or unk or unks, validat*) that were non-specific and could not be adequately refined as determined through testing. Further, Mylan added a "within 50" proximity limiter to ensure that the primary term was reasonably related to a modifier term in order to produce a "hit." This process resulted in amended search terms that were both more precise and yielded a lower volume of hits.

With regard to the QA-Testing terms, the following chart reflects the results of Mylan's effort:

1		Hits	Gigabytes
2	("anion-exchange" OR anionexchange OR "anion exchange") /50 (<drug name> OR <solvents>)	1,003	2.52
3	("ion-exchange" OR ionexchange OR "ion exchange") /50 (<drug name> OR <solvents>)	1,838	5.42
4	*PCRC* /50 (<drug name> OR <solvents>)	76	0.57
5	ALS /50 (<drug name> OR <solvents>)	2,285	7.31
6	bioequiv* /50 (<drug name> OR <solvents>)	17,221	26.12
7	(CAPA or corrective pre/3 "preventive action") /50 <facility name> /50 (<drug name> OR <solvents> OR QA OR "quality assurance" OR QC OR "quality control" OR Chromato* OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 spike" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	82,069	304.06
8	Chromato* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 spike" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	191,390	501.35
9	contami* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 spike" OR "unknown /3 impurit*")	98,309	414.03
10	degradant* or impurit* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 spike" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	513,437	1,020.00
11	detect* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 spike" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	180,412	595.55
12	EI-MS or EL-MS or "electron ionization" /50 (<drug name> OR <solvents>)	17	0.02
13	elude* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 spike" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	286	0.07
14	fatal* /50 (<drug name> OR <solvents>)	3,395	7.26
15	GC or "GC-FID" or GCMS or "GC-MS" /50 (<drug name> OR <solvents>)	145,607	286.75
16	hazard* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 spike" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	26,688	143.63
17	heat /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 spike" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	49,940	121.79
18	HPLC-UV /50 (<drug name> OR <solvents>)	0	0.00
19	HS /50 (<drug name> OR <solvents>)	13,881	36.59
20	incomplet* /5 data* /50 (<drug name> OR <solvents> OR facilities names OR test OR chromatog* OR peak)	10,686	38.72
21	LCMS (<drug name> OR <solvents>)	13,514	16.24
22	mass spectro* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 spike" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	7,746	27.41
23	"method /5 qualification" /50 (<drug name> OR <solvents>)	1,502	2.15
24	noise* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 spike" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	18,527	42.78
25	"noti* w/5 pharm*" /50 (<drug name> OR <solvents>)	5,532	4.57
26	obscure* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 spike" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	740	2.16
27	"press /4 release" /50 (<drug name> OR <solvents>)	6,627	48.28
28	"preventative /5 action" /50 (<drug name> OR <solvents>)	1,025	1.71
29	puri* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 spike" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	226,078	554.80
30	recall* /50 <drug name>	39,471	252.16
31	"repe* /4 error*" /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 spike" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	1,125	4.31
32	residu* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 spike" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	199,202	566.57
33	retrospec* /50 (<drug name> OR <solvents>)	6,054	12.28
34	risk* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 spike" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	210,727	664.78
35	signal* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	20,007	54.34
36	spectro* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	40,174	236.21
37	spike* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	25,380	84.60
38	suitab* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	134,529	387.07
39	toxic* /50 (<drug name> OR <solvents> OR aberran /3 peak* OR abnorm* OR complain* OR "unknown /3 peak" OR "unknown /3 impurit*" OR "unknown /3 contamin*")	44,403	155.75

As illustrated by the QA-Testing terms, although the volume of hits captured by the Revised Search Terms is far less than the overbroad results generated by the Original Search Terms, the Revised Search Terms would still require Mylan to review ***hundreds of thousands*** of emails, ***plus*** attachments to those emails for the QA-testing category alone. Plaintiffs may argue that the modified hit totals are insubstantial, however, Plaintiffs do not bear the burden of processing, storing, and reviewing all of these documents. The Revised Search Terms would still require Mylan and the other Manufacturer Defendants to engage in ***extensive*** efforts to produce responsive documents, but the refined terms would greatly reduce the waste of resources involved in that process.

Notably, QA-Testing is just one of nine (9) categories of primary terms, all of which are equally problematic in terms of volume. (*Compare* Ex. F, Mylan Hit Counts for Revised Search Terms, *with* Ex. G, Mylan Hit Counts for Original Search Terms.) Further, because Mylan's testing of the terms contained in the "authors and inspectors" tab produced such a high volume of false positives based on initial testing, Mylan proposed that these terns should be removed, albeit without prejudice to Plaintiffs' ability to revisit a subset of those terms at a later date after Plaintiffs have had the opportunity to review Mylan's custodial email productions, and are better equipped to determine which authors or inspectors might bear some relevance to Mylan.

The other Manufacturer Defendants have found the search terms to be equally burdensome when applied to their respective data sets. For example, when run against the ZHP Parties'[5] custodial data that has been processed, de-duplicated, and threaded to date,[6] application of the Original Search Terms yields approximately two-thirds (67%) of the custodial data when document families are included, even though these custodians are generally involved with the manufacture of all ZHP API or finished dose products. More specifically, 249,699 documents out of a total of 374,074 documents collected so far for the eight Prinston Pharmaceutical Inc. ("Prinston") custodians would need to be reviewed for responsiveness. The ZHP Parties have so far collected only 186.5 GB of data for the Prinston custodians. If we assume the same number of documents per GB will be returned for the 7,244.16 GB of data collected so far in China, the ZHP Parties could be expected to review approximately 9.7 million documents in addition to the 249,699 Prinston documents.[7]

---

[5] The "ZHP Parties" refers to Zhejiang Huahai Pharmaceutical Co, Ltd., Solco Healthcare U.S., LLC, Prinston Pharmaceutical Inc., and Huahai U.S., Inc.

[6] While the data collection process for the ZHP Parties was delayed due to the travel restrictions and business closure orders associated with the COVID-19 pandemic, the ZHP Parties have made significant progress in collecting data in both China and the U.S. To date, the ZHP parties have collected approximately 7.5 terabytes of data. The ZHP parties have been able to process, de-duplicate, and thread approximately a quarter of a terabyte of data that they have collected from their U.S.-based custodians, and approximately 7.25 terabytes of data that they have collected from their China-based custodians is still being processed.

[7] Because this figure is only an extrapolation from the known data, the ZHP Parties will need to continue to meet and confer with Plaintiffs as they gather actual hit counts and document volumes.

Similarly, use of the search terms as originally drafted, even after implementing modifications agreed to by Plaintiffs and remediation efforts identified by the Teva Defendants' vendor, would require the Teva Defendants to process and review an estimated 4 terabytes of data, and between 5 and six 6 million unique documents, respectively, if applied to all thirty-six (36) custodians.

Accordingly, there can be no meaningful dispute that the Original Search Terms are overbroad. Indeed, the sheer volume of hits renders it facially apparent that the terms are not resulting in "a reasonable search for responsive information pursuant to a reasonably comprehensive search strategy." *Enslin*, 2016 WL 7042206, at *3 (internal quotation and citation omitted). If discovery is to proceed in an efficient and proportionate manner, therefore, the Original Search Terms should not be left undisturbed.

B. Document sampling confirms that the Original Search Terms generate a disproportionate number of false positives.

Given the remarkably high volume of total hit counts returned by the Original Search Terms, it is reasonable to conclude that a large percentage of those hits are false positives. Indeed, hit counts were the parties' agreed-upon baseline metric to be used to determine whether the Original Search Terms needed to be modified, with other testing to follow in some instances. In addition to providing hit counts, the Manufacturer Defendants have also conducted testing and relevancy sampling, thereby generating further proof of overbreadth.

Mylan, for example, selected twelve (12) exemplar Primary Search Terms, taken from each of the challenged categories in the Original ESI Search Terms. As a baseline criterion, all of the exemplar Primary Search Terms tested by Mylan yielded more than 100,000 hits. Mylan then collected sample sets of Microsoft Outlook email files for one-week periods, selected at random, across the relevant time period (for Mylan, January 2011 – the present), across all fifty-two (52) of its ESI custodians. The documents were then processed by Mylan's eDiscovery vendor in accordance with the ESI Protocol. Sample review sets of 100 randomly selected documents were then batched and reviewed by Mylan's counsel for responsiveness to Plaintiffs' Rule 34 Requests.

All told, ninety-four percent (94%) of the documents reviewed by Mylan's counsel were non-responsive. Moreover, less than two percent (2%) of the responsive documents were "true" hits. The remainder were "incidental" hits, meaning that although the documents were potentially responsive to one or more of Plaintiffs' Rule 34 requests, their responsiveness had nothing to do with the primary search term that captured them. This sampling further demonstrates the overbreadth of the Original Search Terms as applied to Mylan's custodial Microsoft Outlook email files due to the extremely low prevalence of responsive documents.

The document sampling conducted by Mylan not only confirmed the overbreadth of the Original Search Terms, but also demonstrated that the Revised Search Terms will significantly reduce Defendants' burden (in terms of time and cost) ***without any apparent prejudice to Plaintiffs***. Specifically, in order to test the efficacy of the Revised Search Terms, Mylan analyzed

whether the responsive documents captured by the Original Search Terms would also be captured by the Revised Search Terms. Mylan determined that the Revised Search Terms captured ninety-three percent (93%) of these responsive documents. Moreover, the seven (7%) of documents not captured by the Revised Search Terms – only six (6) total documents – consisted of SOPs and formal FDA correspondence. Significantly, these are discrete categories of documents that do not relate directly to valsartan, but which Mylan has or will produce from non-custodial sources in response to Plaintiffs' Rule 34 requests. Accordingly, ***none*** of the responsive documents that were captured by the Original ESI Search Terms as identified in Mylan's sample set review would be omitted from production if the Revised Search Terms were adopted by the Court.

Against this mounting evidence of overbreadth, Plaintiffs tried to attack Mylan's methodology, asserting that the sample sizes were not "statistically significant." Dkt. 430-1, 5/11/20 Letter from A. Slater to C. Trischler. More specifically, Plaintiffs' paid consultant opined that, in order to constitute a "statistically significant" sample size, Mylan would need to reach a confidence level of ninety percent (90%) with a five percent (5%) margin of error. Plaintiffs' consultant also assumed a response distribution of fifty percent (50%), which is assuredly too high given the low prevalence of responsive documents identified by Mylan's initial sample review. Nevertheless, based on the existing dataset, Mylan has gone back and done the work suggested by Plaintiffs, including applying Plaintiffs' suggested response distribution of fifty percent (50%). The results of that review are summarized in the following chart.

Primary Term	Total Hits	Time Period	Total Docs Captured (one week)	Sample Set Reviewed	Responsive	True Hit	Incidental Hit
Blood	108,332	February 2013	316	271	**0**	**0**	**0**
cGMP or (current pre/5 manufacturing) or GMP*	784,289	July 2012	393	271	**30**	**25**	**5**
Chromato*	522,652	December 2018	1,735	271	**47**	**16**	**31**
Delete/remove/ trash/shred	1,308,948	October 2017	2,116	271	**6**	**0**	**6**
Detect	517,464	January 2016	1,662	271	**7**	**2**	**5**
Error	1,051,498	May 2018	1,504	271	**15**	**0**	**15**
Fail	782,683	April 2011	586	271	**9**	**0**	**9**
Patrice Hall	104,651	September 2013	165	165	**14**	**0**	**14**
Inspect	682,542	August 2019	1,366	271	**26**	**7**	**19**
Investigat*	780,241	June 2014	2,509	271	**13**	**13**	**0**
Recycl*	890,501	April 2017	5,386	271	**0**	**0**	**0**
Dipesh Shah	240,952	November 2016	342	271	**4**	**0**	**4**
			TOTAL	3,146	**171**	**63**	**108**

Crucially, Mylan's expanded sample review – using a "statistically significant" data set, per Plaintiffs' consultant – resulted in a nearly identical ratio of both non-responsive documents (94%) and "true" hits (2%). This level of consistency after an additional 171 documents for each term were reviewed[8] is notable, and confirms the accuracy of Mylan's initial sample review.

While the ZHP Parties' information systems do not allow it to conduct the testing and sampling conducted by Mylan, the information the ZHP Parties have collected, processed, and de-duplicated to date indicates that approximately 83 percent (i.e., 206,915 out of 249,699) of the documents that would have to be reviewed using the Original Search Terms on the data of the Prinston custodians do not contain the terms most likely to return responsive documents, namely *sartan*, *ndma*, *nitrosamine*, *diovan*, and *exforge*.[9] This means that more than four out of every five documents that Plaintiffs would have to review are highly unlikely to contain responsive information.

Counsel for the Teva Defendants reviewed a random sample of 1200 total documents – 600 of which were returned as positive hits utilizing the original search terms and 600 of which were returned as positive hits utilizing the counterproposal. (Ex. E, Teva Decl. ¶ 9.) Reviewing counsel did not know whether a specific document being reviewed had been identified utilizing the counterproposal search terms or the original search terms. (*Id.*). The sampled review found 8 responsive documents out of 600 utilizing the original search terms (1.3% responsiveness rate) and 89 responsive documents out of 6x00 utilizing the counterproposal search terms (14.8% responsiveness rate). (*Id.*) This disparity suggests not only that the original search terms are plainly overbroad, but that use of the counterproposal search terms, while still including a huge number of irrelevant documents, will result in capture of the truly responsive documents in Defendants' custodial files.

C. The cost of processing, storing, and reviewing documents captured by the Original Search Terms is exorbitant and is not proportionate to the needs of this litigation.

Each Manufacturer Defendant has retained the services of a separate eDiscovery vendor to assist with the processing, production, and storage of ESI. As a result, the specific costs that will be incurred by each Manufacturer Defendant will vary. These differences notwithstanding,

---

[8] The existing dataset did not allow for the review of 271 documents for the primary term "Patrice Hall."

[9] The ZHP Parties expect to address this issue further with Plaintiffs during their ongoing meet and confer process if allowed to do so as planned. During that process, the ZHP Parties will be able to conduct the type of testing and sampling that Mylan has been able to conduct, but because the ZHP Parties' data will have been collected, ***such testing, sampling, and further negotiation with Plaintiffs will not delay the ZHP Parties' review and production of documents. The testing, sampling, and negotiation will be concurrent with the document review.*** Further, although not all of the ZHP Parties' data has been processed, the review of custodial data collected from Prinston is scheduled to begin this week, and counsel for the ZHP Parties has engaged Plaintiffs' counsel in the meet and confer process based specifically on recent analyses of Prinston's data.

Defendants certainly recognize that ESI discovery in an MDL will involve significant expense. Due to the overbreadth of the Original Search Terms, however, the Manufacturer Defendants face runaway costs that are disproportionate to the needs of this litigation.

By way of illustration, Mylan has consulted with its eDiscovery service provider, Consilio, and has determined that Mylan will incur the following costs to process, store, and review custodial Microsoft Outlook email files captured by the Original Search Terms, and the Revised Search Terms:

- Original Search Terms: **$7,865,000**
- Revised Search Terms: **$3,375,000**

(Ex. C, Consilio Decl. ¶¶ 6-7.) Notably, these figures do not include estimates related to processing, storage, and review of either non-custodial documents, or non-email custodial documents.

The contrast between these estimated costs is stark. Moreover, while Mylan's costs associated with processing, storing, and reviewing documents captured by the Revised Search Terms are substantial, they represent a savings of well over fifty percent (50%) compared to what Mylan would incur by using the Original Search Terms—again, with no apparent downside to Plaintiffs.[10]

The costs to be incurred by the other Manufacturer Defendants are also extraordinarily high. For example, as previously noted, the ZHP Parties project that the application of the Original Search terms could yield approximately 10 million documents, inclusive of families. At a review rate of 300 documents per reviewer per day (which includes responsiveness, confidentiality, and privilege review), this task could require the ZHP Parties to employ at least 220 document reviewers to complete the review by end of November, meaning the review alone would cost several million dollars.[11]

The Teva Defendants requested an estimate from their vendor as to the review cost associated with utilizing the original search terms. (Ex. I, ULX Decl. ¶ 2.) The Teva Defendants'

[10] All of the foregoing cost estimates related to Mylan's data are based upon conservative assumptions for reduction in review volume based upon de-duplication and email threading. (Ex. C, Consilio Decl. ¶ 8.) These assumptions reflect industry norms, with a weighted bias based on Mylan's typical data behavior. (*Id.* ¶ 9.)

[11] As the ZHP Parties are only able to provide projections at this time, they will need to continue to meet and confer with Plaintiffs as they gather actual document counts and resulting review costs. Plaintiffs should not be permitted to avoid their meet and confer obligation with the ZHP Parties simply because, due to COVID-19, the ZHP Parties have not yet been able to pull this information from the 7.25 terabytes of information they have only just been able to collect in the past few weeks.

vendor estimated and compared the anticipated post-deduplication cost of utilizing the original search terms as opposed to Mylan's counterproposal:

- Original Search Terms: **$3,708,500**
- Revised Search Terms: **$1,976,300**

(*Id.* ¶ 3.)[12] Note that this number is simply an estimate for reviewing the custodial documents of the 16-custodian sample set. Applied across all 36 of the Teva Defendants custodians these costs are likely to double and result in nearly $3.5 Million in additional costs, before even accounting for increased processing and hosting costs.

III. Cost-shifting is appropriate if the Defendants are required to use the Original Search Terms.

To the extent the Court declines to adopt the Revised Search Terms, Plaintiffs should be required to share in the costs of ESI discovery.

Under the Federal Rules of Civil Procedure, "the presumption is that the responding party must bear the expense of complying with discovery requests." *Oppenheimer Fund, Inc. v. Sanders*, 437 U.S. 340, 358 (1978). Yet, a producing party "may invoke the district court's discretion under Rule 26(c) to grant orders protecting him from 'undue burden or expense' in doing so, including orders conditioning discovery on the requesting party's payment of the costs of discovery." *Id.* A party may therefore move to condition discovery on the other party's payment of the discovery costs. *Race Tires America, Inc. v. Hoosier Racing Tire Corp.*, 674 F.3d 158, 171 (3d Cir. 2012).

In a procedurally similar case, *Boeynaems v. LA Fitness International, LLC*, the district court granted the defendants' request to shift costs where the plaintiffs demanded extensive and burdensome discovery while class certification was pending, and after defendants had already produced a number of documents:

> [T]he Court mandates cost allocation as fair and appropriate. The Court concludes that where (1) class certification is pending, and (2) the plaintiffs have asked for very extensive discovery, compliance with which will be very expensive, that absent compelling equitable circumstances to the contrary, the plaintiffs should pay for the discovery they seek. If the plaintiffs have confidence in their contention that the Court should certify the class, then the plaintiffs should have no objection to making an investment. Where the burden of discovery expense is almost entirely on the defendant, principally because the plaintiffs seek class certification, then the plaintiffs should share the costs.

[12] This estimate includes the following standard assumptions: (1) a typical 30% responsive rate across the full review set; (2) a 30% redaction rate across the review set to account for potential Personal Information, HIPAA, Non-Responsive/Other Teva product redactions, and privilege redactions; and (3) a 25% privilege rate across the responsive production set encompassing privilege sweep, QC, and consistency review. (Ex. I, ULX Decl. ¶ 3.)

285 F.R.D. 331, 341 (E.D. Pa. 2012). The court further observed that "[s]hifting the cost burdens of discovery, both for ESI and paper discovery, is no longer rare." *Id.* at 336; *see also Schweinfurth v. Motorola, Inc.*, No. 1:05-cv-24 2008 WL 4449081 (N.D. Ohio Sept. 30, 2008) (ordering plaintiffs to pay half of discovery costs because plaintiffs' requested discovery, though potentially relevant, was massive in scope and would not necessarily lead to the discovery of evidence admissible at trial); *Universal Delaware, Inc. v. Comdata Corp.*, No. 2:07-cv-1078, 2010 WL 1381225, at *8 (E.D. Pa. Mar. 31, 2010) (granting cost sharing motion).

This Court has similarly endorsed discovery cost-shifting under appropriate circumstances. *See Constantino v. Atlantic City* 152 F. Supp. 3d 311, 337 (D.N.J. 2015) (encouraging the parties to come up with an efficient discovery plan, including cost-shifting); *Major Tours, Inc. v. Colorel*, No. 1:05-cv-3091, 2009 WL 3446761, at *6 (D.N.J. Oct. 20, 2009) (allocating discovery costs and noting that relief is appropriate "when electronic discovery imposes an undue burden or expense on the responding party").

Cost shifting is appropriate here because, as demonstrated by the data presented above, the use of the Original Search Terms will impose in a significant financial burden on the Manufacturer Defendants. Because the Plaintiffs have refused to meaningfully meet and confer regarding modifications to the Original Search Terms that would reduce that burden, Plaintiffs should be required to shoulder a portion of the cost for all of the unnecessary processing, storage, and attorney review that will result. Therefore, to the extent this Court declines to adopt the Revised Search Terms, Defendants request the entry of an order allocating to Plaintiffs the costs of processing, reviewing, and storing the ESI they demand.

IV. Conclusion.

For all of the foregoing reasons, the Manufacturing Defendants respectfully request that this Court modify its December 23, 2019 Order and adopt the Revised Search Terms in their entirety.

Respectfully submitted,

Clem C. Trischler

Enclosures

c: All Counsel of Record (via ECF service)